DOCUMENT RESUME

ED 158 223 CS 004 224

AUTHOR Weiner, Cheryl J.
TITLE The Effect of Training in Questioning and Student Question Generation on Reading Achievement.
PUB DATE Mar 78
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (Toronto, Canada, March 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Cognitive Processes; Elementary School Students; Intermediate Grades; *Questioning Techniques; *Reading Achievement; *Reading Comprehension; *Reading Research; *Training Techniques

ABSTRACT

Based on the assumption that the active involvement of the reader is required to give meaning to the printed word, a study was devised to test whether reading comprehension can be increased if students are taught to formulate questions about what they read. The study had three phases: a pilot study during which fifth grade top readers were observed, a training program sequence consisting of activities that directed students in formulating generic questions about what they read, and an instructional unit on ecology that was divided into nine lessons. Participants in the second and third phases consisted of 91 sixth graders working in 27 dyads and divided into four treatment groups. Two groups were trained in questioning and all four groups participated in the instructional units. Reading achievement in all four groups was measured by a multiple choice test and a short answer essay test. Results of the study indicated that the students in the two groups that received training in questioning formulation scored significantly higher on both the multiple choice test and the essay test than those in the two groups that had not been trained. There was no significant difference between the students who formulated questions during the instructional sequence and those who did not. (MAT)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED158223

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE EFFECT OF TRAINING IN QUESTIONING AND STUDENT QUESTION GENERATION ON READING ACHIEVEMENT

CHERYL J. WEINER

Educational Testing Service
Princeton, NJ 08541

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Cheryl J. Weiner

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Presented at the annual meeting of the
American Educational Research Association
Toronto, March 1978

CS004224

Printed in USA

This study emerged from a theoretical framework which asserts that active processing in both reading and learning facilitates achievement. Self-generated questioning was posited as an active cognitive strategy that could be used while reading to increase comprehension. Training students to generate specific questions related to increasing their reading achievement was the focus of the study.

Reading models proposed by Gibson (1975), Goodman (1970), and Smith (1971) assert that active involvement in reading is needed for meaning to be created. Inducing students to ask questions about what they read so that they can actively interact with a text constitutes one of the major components of various reading study skills methods (Robinson, 1946; Thomas & Robinson, 1972). Although these methods have been referred to extensively in major textbooks on reading instruction as effective in eliciting achievement gains in learning from prose, there is limited research data to support this claim (Diggs, 1975; Donald, 1967). Moreover, the self-generated questioning component of these study methods has not been separately validated as a significant variable contributing to achievement. Nor has there been an operational definition given for the types of questions that should be used. Other reading programs that use student generated questions are related to individual tutoring and disabled readers (Manzo, 1970) and pre-reading activities (Stauffer, 1969), and have similar limitations. Consequently, although there is a sizable literature that suggests that self-generated questioning is an important reading skill, there is not a firm research base that justifies this position.

Support for active learning comes from Piaget's theory of cognitive development which is based on the tenet that learning takes place through an active reorganization of data (Ginsburg & Opper, 1969). Ausubel's theory of

verbal learning (1963) and the mathemagenic theory of prose learning (Rothkopf, 1970) extend the concept of active processing to reading prose texts. In mathemagenic studies, adjunct question aids have been manipulated by the experimenter in order to effect internal structuring which in turn has facilitated learning (Faw & Waller, 1976; Frase, 1970).

Two sets of researchers have investigated the effect on achievement of having students generate questions while reading rather than having the experimenter manipulate questions. Within the mathemagenic context, in two studies, Frase and Schwartz (1975) found that asking students to generate questions while they read increased their recall of information. Schermerhorn, Goldschmid, and Shore (1975) also demonstrated that students generating questions about prose facilitated their learning from it. Both of these sets of studies used student dyads effectively as instructional groupings.

However, the results of these three studies also have some important limitations which need to be addressed by further research. First, Schermerhorn, et al. (1975) did not separate out the effects of reading from question generation by having a control group read the material without engaging in questioning. Second, in Frase and Schwartz (1975) the instructions given to students severely limited the types of questions that they would ask. Students had to identify the lines that contained the answers to their questions. This led to the creation of severely restricted questions. Third, both studies used a single achievement measure which measured only literal responses. Fourth, neither study trained students to ask questions nor gave them a chance to practice their questioning strategy over a long period of time.

This study addressed itself to the limitations in both the reading research cited and these three questioning studies. The purpose of this study was to investigate the effects of student question generation on reading achievement.

The study was designed to answer the following questions.

1) How does training students to generate questions about what they read affect reading achievement?
2) How does inducing students to generate questions about what they read during an instructional unit affect their reading achievement?
3) What is the nature of the interaction effect between training and questioning on reading achievement?
4) What is the effect of other training and questioning relationships on reading achievement?

METHOD

Subjects:

All subjects came from one public elementary school in Springfield, Oregon. Sixteen fifth-grade students who were members of three different classes but were in their top reading groups were used in a pilot study. All ninety-one students in three in-tact sixth-grade classes participated in the experimental study. In each of the three sixth-grade classrooms students were randomly divided into four treatment groups. They were paired into dyads by the experimenter. Then each of the four groups in each class were randomly assigned to be in one of the four treatment groups: Trained-Questioners (TQ), Trained Non-Questioners (TNQ), Untrained Questioners (UQ), and Untrained Non-Questioners (UNQ).

Materials:

There were two parts to the study: the training program sequence and the instructional unit sequence. The training program was developed by the experimenter. The curriculum consisted of a series of worksheets and activities which utilized direct instruction techniques to train students to generate six generic questions about what they read. These were:

1) How does this passage or chapter relate to things that I already know about?
2) What is the main idea of this passage or chapter?
3) What are five important ideas that the author develops that relate to the main idea?
4) How does the author put the ideas in order?
5) What are the key vocabulary words? Do I know what they all mean?
6) What special things does the passage make me think about?

Each lesson concentrated on program goals. Lessons included instructional activities and opportunities for practice or review. Both oral and written drills were used. A mastery post-test was given after training to measure whether students could generate the six questions.

The training program was designed for use with any expository reading material divided into chapters or sections. For this study, the school district requested that students read the book, Drugs and People (Read, 1969), so the program was superimposed over that text.

The curriculum used in the instructional unit sequence on ecology was developed by the staff of the Effective Teacher Education Program at the Far West Laboratory for Educational Research and Development as part of a large-scale study on teacher-directed questioning and its effect on student achievement and attitudes (Gall, Ward, Berliner, Cahen, Crown, Elashoff, Stanton, &

Winne, 1976). The curriculum materials used consisted of a set of 10 printed handouts written at the sixth-grade level, which were developed at the Laboratory, and 3 commercial audio-visual aids: a film and two film strips. The instructional unit sequence was divided into 9 lessons.

Assessment Instruments:

Reading achievement was measured by a multiple-choice test and a short answer test. They were both constructed by The Far West Laboratory staff in conjunction with their instructional materials (Gall & Crown, 1974). One form of each test was used as both the pre- and post-test instruments. These tests had already been used by The Laboratory with sixth-grade students, and their reliability and content validity were considered to be acceptable.

Content validity for both tests was checked by judging whether the items on both tests could be answered by the instructional materials alone.

In order to assure the reliability of the multiple-choice test for this sample, a local reliability estimate was computed using the Kuder Richardson 20 formula. The pre-test and post-test reliability were .30 and .76, respectively. The criteria for scoring the short answer tests which are listed in Gall and Crown (1974) were modified for use. Two scorers read each test and had an inter-rater reliability of .95 on both the pre- and post-test scoring using the Pearson Product Moment Correlation Coefficient.

Procedures:

The study had three phases: a pilot study, a training program sequence, and an instructional unit sequence. The 3 day pilot study was used to make decisions about how to implement both the training program and the instructional unit. During the 9 day training program sequence, the TQ and TNQ groups were trained to ask questions about what they read. During the 11 day

instructional unit sequence on ecology, students in all four treatment groups read a passage, completed a written worksheet, and interacted with a partner in dyads. The worksheets given to the TQ and UQ groups directed them to generate six questions about the reading. The worksheets given to the TNQ and UNQ groups directed them to complete various language arts activities related to the content of the ecology unit.

RESULTS

Reading achievement was measured by a multiple-choice test called the Ecology Information Test and by a short answer test call the Essay Test. Both were given to all four treatment groups before and after the instructional unit sequence.

A 2 (training) x 2 (questioning) factorial design was used to investigate the two main effects of training and questioning on reading achievement. A two-way Analysis of Covariance was used to analyze the post-test scores on the two tests. The pre-tests were the covariates. The mean score of the dyad of partners was the unit of analysis. There were 27 dyads analyzed: 8 in the TQ group, 5 in the UQ group, 7 in the TNQ group, and 7 in the UNQ group. Mulitple planned comparisons were computed using the Scheffe test. Statistical significance was at the .05 level.

In analyzing both the Ecology Information Test and the Essay Test scores, the means and standard deviations for the pre- and post-tests indicated that the post-test differences among treatment groups were not the result of existing pre-test differences.

On the Ecology information Test, the post test groups were ranked from high to low: TNQ = 13.71 > TQ = 12.00 > UQ = 11.67 > UNQ = 10.00. The UQ

group had the largest standard deviation at 5.42. Adjusting the mean scores for the Analysis of Covariance had two effects. First it changed the rank order of the groups to: TNQ = 14.527 > UQ = 11.69 > TQ = 11.37 > UNQ = 9.86. Second, the adjustment of scores resulted in a disordinal interaction between training and questioning. The two main effects were not influenced by the adjustment.

On the Essay Test the post-test groups were ranked from high to low: TQ = 14.34 > TNQ = 12.78 > UQ = 11.50 > UNQ = 11.21. The standard deviation for the UQ group was slightly higher than the others at 4.15. The adjustment of the post-test mean scores changed the rank ordering to: TQ = 14.14 > TNQ = 12.45 > UNQ = 12.16 > UQ = 11.10. This did not influence the main effects, but did change the interaction between training and questioning from ordinal to disordinal.

In comparing the two sets of data from the Ecology Information Test and the Essay Test, a training effect was evident on both achievement tests. The direction of the effect was positive. The questioning effect was minimal in both tests. The interaction between questioning and training demonstrated an interesting reversal. For the Ecology Information Test, the TNQ had the highest mean score at 14.52, whereas the UNQ had the lowest at 9.86. On the Essay Test, the TQ had the highest mean score at 14.14, while the UQ had the lowest at 11.10.

The only statistically significant results came from the scores on the Ecology Information Test. The two-way ANOCOVA showed a significant F-ratio at the .05 level of 4.64 for the main effect of training, and 4.99 for the interaction between training and questioning. The F-ratio for the questioning main effect was 0.47 and was not significant. Table 1 summarizes the ANOCOVA results for the Ecology Information Test.

Table 1

Summary of Two-Way Analysis of Covariance on the Ecology Information Post-Test with Scores on the Pre-Test Used as the Covariate

Source	Degrees of Freedom	Sum of Squares	Mean Sum of Squares	F Value
Training	1	34.03	34.03	4.64*
Questioning	1	3.47	3.47	0.47
Training x Questioning	1	36.55	36.55	4.99*
Within	22	161.22	7.33	
Total	26	249.91		

*$p < .05$

Scores on the Essay Test did not achieve statistical significance, they did show trends in the same direction as the main effect. However, the interaction effect was ordinal. For the training main effect, the F-ratio was 1.92; while it was .20 for the questioning main effect. Table 2 summarizes the ANOCOVA results for the Essay Test.

The Scheffé Test, analyzing planned comparison on the Ecology Information Test, indicated that there were two statistically significant comparisons at the .05 level. The Trained Non-Questioners group at 14.52 was significantly higher than the Trained Questioners group at 11.37. Also, the combined mean scores of the Trained Questioners at 11.37, the Untrained Questioners at

Table 2

Summary of Two-Way Analysis of Covariance on the Essay Test with Scores on the Pre-Test Used as the Covariate

Source	Degrees of Freedom	Sums of Squares	Mean Sums of Squares	F Ratio
Training	1	16.56	16.56	1.92
Questioning	1	1.72	1.72	0.20
Training X Questioning	1	11.77	11.77	1.36
Within	22	190.16	8.64	
Total	26	335.78		

$p < .05$

11.69, and the Trained Non-Questioners at 14.52, were significantly higher than the scores of the Untrained Non-Questioners at 10.76, which served as the control group. No comparisons made on the Essay Test scores were significant.

Predictions of individual scores on the Ecology Information Test would vary for students in the four treatment groups. For example, if a student were in the TQ group, his adjusted score would range between 9.06 and 13.66. If a student were in the TNQ group, his score would range between 11.97 and 17.07. A sizeable overlap exists between the scores in both groups, but the scores in the TNQ group were two points below. A gap of this size demonstrates the power of the statistical differences between these two groups. There were no other two groups which were significantly different from each

other on either test. Any other predictions would have to be based on the total mean for each test.

Based on the results for the Ecology Information Test, it is possible to answer the three research questions related to the two main effects and their interaction. The students in the two groups which received training in question generation scored significantly higher on the Ecology Information Test than the students in the two groups which were not trained. On the other hand, there was no significant difference between the scores of students who generated questions during the instructional unit sequence and those who did not. By combining questioning and training, however, there was a significant disordinal interaction. Students who were trained but did not question scored higher than students who were trained and did question. Students without training who questioned scored higher than students without training who did not question. Thus, the interaction of the two effects which resulted in the high scores of the TNQ group contributed to the training effect.

The planned comparison results address the fourth research question, which concerned the relationship between the treatment groups. Since the TNQ group scored higher than the TQ group on the Ecology Information Test, getting students to practice what they had learned did not necessarily help them remember literal information better than training them to mastery in questioning and having them do other language arts activites during instruction. However, the combined treatment groups' superiority over the UNQ group indicated that just reading the passage and not receiving treatment on either questioning or training was not as effective in increasing students' achievement on tests related to what they had read. No other relationships were found.

DISCUSSION

There were four methodological factors which could have influenced the results. First, the students who participated in the training program received four hours of specialized instruction from the experimenter. However, the untrained students received specialized instruction from their teachers during that four hours. Also, the experimenter interacted with the untrained students enough prior to the instructional unit sequence so that any familiarity affects should have been minimized.

The second factor related to the use of the dyad. Dyadic interaction represented a valid instructional mode used in other studies (Frase & Schwartz, 1975, Schermerhorn, et. al., 1975). However, the sustained dyadic interaction necessary for this study proved to be problematic. To mitigate against the effects of absenteeism and unusual behavior within dyads, rigorous criteria for the inclusion of a dyad in the data analysis were established. In addition, the dyad was chosen as the unit of statistical analysis in order to assure the most conservative results (Peckham, Glass, & Hopkins, 1969).

The third factor concerned the validity of the training program. The particular training questions, their sequencing, and the instructional strategies used to teach them had not been separately analyzed. However, mastery of the questions was checked at the end of the training program sequence and the instructional unit sequence. The question sequencing and instructional strategies used were tested in the pilot study and reflected standard practices in reading and direct instruction lessons. In general, students learned the questions and applied that learning appropriately throughout the training program sequence.

The fourth factor was the logistical constraint of working in a public school during regular school hours. Time was lost. Lessons were interrupted. Students were absent. Procedures had to be changed to accommodate these unpredictable events. Nevertheless, these difficulties are a basic component of public school instruction. The significant results indicated that the treatments were robust enough to weather alternations and reflected findings that could be generalized to classroom activites.

Assuming that these factors did not influence the results in a major way, four findings emerged which have implications for both classroom teachers and researchers. First, training students to engage in a covert but active cognitive strategy while reading, such as generating questions, may be an effective way to improve students' achievement in reading. This finding can be supported by the generative, mathemagenic, and meaningful prose learning theories, as well as the reconstructive theories of reading comprehension. In addition, these results confirm the worth of having students ask questions about their reading which is part of the SQ3R process, a standard reading instructional technique.

This study is differentiated from other question generation studies because it related questioning to two achievement measures which were specific to the reading covered but which were general tests of achievement. They were not directly related to the types of questions generated by students. The general nature of the achievement gains suggest that this particular training program in questioning has a high level of transfer. Training did not just affect inquiry skills or problem-solving skills. Rather, training generalized to typical objective test achievement measures.

Second, making students manifest the covert strategy of questioning in overt observable behaviors, such as completing worksheets which direct them to question, may interfere with their reading achievement. This finding calls into question the utility of drill work once a skill has been mastered. Also, this indicates that some direct instruction activites which emphasize overlearning may be counter-productive for some students. The efficacy of rote worksheet drills as opposed to more open ended language arts worksheet activites should be further explored.

Third, achievement tests may measure different types of cognitive processes. Given different results on the multiple-choice and open-ended measures in this study, more than one type of measure was warranted. The long term effect of training and question generation treatments on multiple measures needs to be investigated further.

Fourth, grouping students into dyads is a viable instructional technique, which deserves more careful consideration as an alternative that could promote more efficient prose learning. However, the effect of dyadic interaction on achievement was not investigated in this study.

The results of this study strongly suggest that training students to generate questions during reading positively affected their reading achievement, but more comprehensive studies are needed to further investigate the effects of teaching cognitive learning strategies to students. Analyzing components of training programs, comparing various types of strategies, and using multiple achievement measures in other areas besides reading constitute future research activities necessary to explicate the findings cited here.

References

Ausubel, D.P. The psychology of meaningful verbal learning. New York: Grune and Stratton, 1963.

Diggs, Y.M. The relative effectiveness of the SQ3R method, a mechanized approach, and a combination method for teaching remedial reading to college freshman. In G.H. McNinch and W.D. Miller (eds.), Reading: Convention and Inquiry. Twenty fourth Yearbook of National Reading Conference. South Carolina: National Reading Conference, 1975, 161.

Donald, M. The SQ3R method in grade seven. Journal of Reading, 1967, 11, 33-35.

Faw, H.W. and Waller, G.T. Mathemagenic behaviors and efficiency in learning from prose materials: Review, critique and recommendations. Review of Educational Research, Fall 1976, 46(4), 691-720.

Frase, L.T. Boundary conditions for mathemagenic behaviors. Review of Educational Research, 1970, 40, 337-348.

Frase, L.T. and Schwartz, B.J. Effect of question production and ansering on prose recall. Journal of Educational Psychology, 1975, 67, 628-635.

Gall, M.D., Ward, B.A., Berliner, D.C., Cahen, L.S., Crown, K.A., Elashoff, J.D., Stanton, G.S., & Winner, P.H. The effects of teacher use of questioning techniques on student achievement and attitudes. Volume 1, Final Report. San Francisco: Far West Laboratory for Educational Research and Development, 1976.

Gall, M.D., Crown, K.A., Madsen, M.L., Thorn, L.I., Ward, B.A. The effects of teacher use of questioning techniques on student achievement and attitudes. Vol. 2: Curriculum and treatment materials. San Francisco: Laboratory for Educational Research and Development, 1974.

Gall, M.D. & Crown, K.A. The effects of teacher use of questioning techniques on student achievement and attitudes. Vol. 3: Achievement tests and attitude scales. San Francisco: Far West Laboratory for Educational Research and Development, 1974.

Gibson, E.J. and Levin, H. The psychology of reading. Cambridge: MIT Press, 1975.

Ginsberg, H. and Opper, S. Piaget's theory of intellectual development: An introduction. Englewood Cliffs, N.J.: Prentice-Hall, 1969.

Goodman, K. Reading: Process and program. National Council of Teachers of English. 1971.

Manzo, A.V. Reading and questioning: The ReQuest procedure. Reading Improvement, 1970, 7, 80-83.

Peckham, P.D., Glass, G.V. & Hopkins, R.D. The experimental unit in statistical analysis. The Journal of Special Education, Winter 1969, 3(4), 337-349.

Read, D.A. Drugs and people. Boston: Allyn and Bacon Inc., 1969.

Robinson, F.P. Effective study. New York: Harper & Bros., 1946.

Rothkopf, E.Z. The concept of mathemagenic activites. Review of Educational Research, 1970, 40, 325-336.

Schermerhorn, S.M., Goldschmid, M.L., and Shore, B.M. Learning basic principles of probability in student dyads: A cross-age comparison. Journal of Educational Psychology, 1975, 67,(4), 551-557.

Smith, F. Understanding reading. New York: Holt, Rinehart, & Winston, 1971.

Stauffer, R.G. Directing reading maturity as a cognitive process. New York: Harper & Row, 1969.

Thomas, E.L. and Robinson, H.A. Improving reading in every class (abridged). Boston: Allyn & Bacon, 1972, 69-140, 134.